

Face-to-diagonal reduction of Kramers-Wannier approximation for cubic lattice particle systems with nearest neighbour exclusion

Igor G. Vladimirov*

UNSW Canberra, ACT 2600, Australia

The paper is concerned with interacting particle systems on the simple cubic lattice obeying the nearest neighbour exclusion (NNE). This constraint forbids any two neighbouring sites of the lattice to be simultaneously occupied, thus reducing the set of admissible configurations for the cubic cell and its subclusters such as edges and faces. This reduction extends applicability of Kikuchi's Cluster Variation Method (CVM) with higher-order clusters to systems with complex site configurations and short-range ordering, which would be impractical beyond the NNE framework because of the "curse of dimensionality". For edges of the cubic cell, which are the operational clusters of the Bethe-Peierls entropy approximation in the CVM hierarchy, the edge-to-site reduction of the entropy cumulants was studied previously. In extending the earlier results, we develop a face-to-diagonal reduction of the Kramers-Wannier entropy approximation of the CVM in the NNE setting. We also outline an application of the resulting approximation to thermodynamic modeling of disordered condensed media, such as liquid silicates, and discuss combinatorial and numerical aspects of the implementation of this approach.

PACS numbers: 82.60.-s, 02.50.Cw, 65.40.Gr, 82.60.Lf.

I. INTRODUCTION

The present study continues the theme of [1] and is concerned with the statistical mechanical modeling of classical interacting particle systems on the simple cubic lattice with nearest neighbour exclusion (NNE). Assuming the site configurational space to be a finite set, and distinguishing one of its elements as *vacancy* while interpreting the others as *occupied* site configurations, NNE forbids any two nearest sites of the lattice to be simultaneously occupied. With the number of site configurations not being limited to two, NNE generalizes a similar constraint which is used in the hard-core lattice liquid/gas models [2, 3]. Moreover, it provides a more flexible setting for the modeling of *disordered* condensed media in comparison with the face-centered-cubic (FCC) lattice formulation.

For NNE-constrained cubic lattice particle systems, we develop a *face-to-diagonal reduction* of the Kramers-Wannier entropy approximation [4]. The latter constitutes the second level in the hierarchy of the Cluster Variation Method (CVM) approximations [5–11] and takes into account the equilibrium statistical correlations in the particle system within faces of cubic cells. Accordingly, the configurational entropy of the system per lattice site, which is known to be intractable in three dimensions, is approximated by a linear combination of the site, edge and face entropies weighted by Kikuchi-Barker coefficients for the simple cubic lattice. Thus, the faces of cubic cells of the lattice play the role of basic clusters with their edges and sites as subclusters.

The NNE constraint allows the Kramers-Wannier entropy approximation to be reduced to the site and face-diagonal entropies, where the latter are associated with face diagonals of cubic cells. More precisely, the resulting entropy approxi-

mation is organized as a linear combination of the site and face-diagonal entropies along with the Shannon mutual information [12, p. 19–22] between the occupancies of the nearest neighbours and the face diagonals, which are induced by the NNE constraint. That is, the entropy approximation is essentially reduced to one- and two-site clusters. The present study takes advantage of this dimensionality reduction potential of the NNE setting in order to advance practical applicability of the Kramers-Wannier approximation towards lattice models of disordered condensed media with complex site configurations and short-range ordering.

We apply the NNE-induced face-to-diagonal reduction of the Kramers-Wannier approximation of the *configurational* entropy to a class of chemical systems, where site configurations represent spatial arrangements of *coordination entities* consisting of a *central atom* at a site of the simple cubic lattice and a surrounding array of ligands. The locations of the central atoms are subjected to NNE, whilst the ligands are allowed to reside at sites of an interstitial lattice and may be shared by several "overlapping" coordination entities. The overlap induces additional geometric constraints which single out a class of admissible pairs of coordination entities centered at the end sites of face diagonals.

The energetics of the system is parameterized by interaction energies which are ascribed to representatives of the isotropy equivalence classes into which the admissible face-diagonal configurations are partitioned by the action of the full octahedral symmetry group of 48 isometries of the simple cubic lattice [13]. We develop a theory which allows the equilibrium Gibbs energy and related thermodynamic quantities to be approximately computed for given values of the energy parameters. This approach is based on minimizing the variational free energy density approximation (per central atom) over admissible face-diagonal and site probability mass functions (PMFs) subject to the compatibility (marginalization) and balance relations.

The constrained minimization problem is solved by using a

*Electronic address: igor.g.vladimirov@gmail.com

separation-of-variables technique, not dissimilar to that in Dynamic Programming. More precisely, the optimization problem is decomposed into a family of pairs of optimization problems which share a common scalar parameter but are solved separately, with the dependencies on the pressure and chemical composition of the system entering these two problems in isolated ways. The solution of the first problem is reduced to finding a root of a decic polynomial, whilst the second problem resembles the Bethe-Peierls approximation for an FCC lattice, though with different coefficients in the entropy cumulants. These solutions are then “assimilated” by minimization over the master parameter, which leads to a complicated (non-additive) dependence on the pressure and chemical composition. This computational approach is applicable to a wide range of pressure values (in principle, including tectonic pressures).

In order to develop the NNE-constrained entropy approximation and the solution of the optimization problem, we employ the probabilistic concepts such as the above mentioned Shannon information and conditional entropy [12, p. 17–20] (which are used here similarly to [1]) together with Boolean random variables and logical operations on them, including the Sheffer stroke [14, p. 51]. This combination provides an efficient machinery for entropy theoretic computations under geometric constraints.

The face-to-diagonal reduction of the Kramers-Wannier approximation is then applied to thermodynamic modeling of a binary liquid silicate $\text{SiO}_2\text{--M}_2\text{O}$ formed from silica and the oxide of a univalent metal M, and the combinatorial and numerical aspects of this application are discussed. Using the structural model from Section IV of [1], which relies on the qualitative insights into the internal structure of silicate melts [15, 16], the liquid silicate is modelled as an assemblage of Si–O–Si, Si–O–M and M–O–M second nearest neighbour bonds (SNNBs) centered at oxygen atoms residing at sites of the simple cubic lattice subject to NNE.

Since the present study takes into account face-diagonal correlations, it also leads to a refined internal energy model which, as mentioned above, sums the energies of interaction between SNNBs at the end sites of face diagonals of cubic cells. This includes, as a particular case, the approach of the existing thermodynamic models to liquid silicates [15, 16], such as the Quasi-Chemical Model and its modifications [17–19], where the internal energy is assumed to be composed of SNNB energies.

Practical model calibration for specific systems, such as sodium silicate $\text{SiO}_2\text{--Na}_2\text{O}$, depends on availability of an efficient numerical algorithm for solving a particular non-convex minimization problem. This requires an additional research into the implementation of the model (see Section XV B for details) and is beyond the scope of the present paper.

The organization of the paper is as follows. Section II describes the NNE-constrained cubic lattice particle systems being considered. These are instantiated in Section III, which specifies the set of occupied site configurations and related balance equations for chemical systems with short-range ordering, such as liquid silicates. Section IV revisits the edge-

to-site entropy reduction from [1] for completeness. Section V expresses the face entropies in terms of the appropriate face-diagonal entropies. Sections VI and VII describe marginalization and other constraints for the face-diagonal PMFs. Section VIII carries out the face-to-diagonal reduction of the Kramers-Wannier entropy density approximation under NNE. In combination with the internal energy model of Section IX, the resulting entropy density estimate is used in Section X in order to formulate the approximate computation of the Gibbs energy through minimizing the variational free energy estimate per central atom over admissible face-diagonal and site PMFs subject to the marginalization and balance constraints. The solution of this constrained optimization problem is considered in Sections XII–XV under an additional isotropy assumption of Section XI which further reduces the problem dimensionality by an order of magnitude. To this end, Section XV develops the separation-of-variables technique based on the conditioned representation of the site and face-diagonal entropies and the internal energy density from Section XII along with isotropic versions of balance and marginalization equations from Sections XIII and XIV. Section XVI outlines an application of the statistical mechanical approach to thermodynamic modeling of a binary liquid silicate. Concluding remarks are given in Section XVII. Appendices provide subsidiary material.

II. NNE-CONSTRAINED CUBIC LATTICE SETTING

We consider an interacting particle system at thermodynamic equilibrium on the simple cubic lattice \mathbb{Z}^3 with a finite *site configurational space*

$$\Omega := \{0\} \sqcup W. \quad (1)$$

Here, 0 is interpreted as *vacancy*, W is a set of configurations for an *occupied* site of the lattice (so that $0 \notin W$), and \sqcup denotes the union of disjoint sets. The equilibrium spatial arrangement of the particle system is described by an Ω -valued *homogeneous* random field $\xi := (\xi_z)_{z \in \mathbb{Z}^3}$, where ξ_z is the state of site $z \in \mathbb{Z}^3$. The homogeneity of ξ is understood in the usual sense as the invariance of its multi-point probability distributions with respect to translations of the lattice [20, 21]. The states of sites of the cubic cell of the lattice

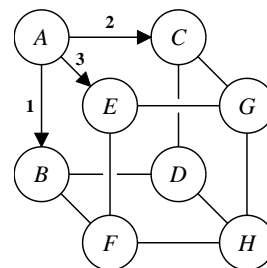


FIG. 1: The states A, \dots, H of sites of the cubic cell. The nearest neighbours are connected by edges. The arrows represent the Cartesian coordinate axes.

are denoted by A, B, C, D, E, F, G, H as shown in Fig. 1. These are identically distributed random variables with values in the set Ω . Their common marginal probability mass function (PMF) $S := (S_u)_{u \in \Omega} \in [0, 1]^\Omega$, which we will refer to as the *site PMF*, is defined by

$$S_u := \mathbf{P}(A = u), \quad u \in \Omega, \quad (2)$$

where $\mathbf{P}(\cdot)$ is the underlying probability measure. Assuming that the equilibrium random field ξ is spatially ergodic, S_u represents the relative fraction of those sites in a macroscopically large (for simplicity, cubic) fragment of the lattice which are at a given configuration $u \in \Omega$:

$$\lim_{N \rightarrow +\infty} \frac{\#\{z \in Q_N : \xi_z = u\}}{N^3} = S_u, \quad (3)$$

where the convergence holds almost surely. Here, $\#(\cdot)$ is the number of elements in a finite set, and

$$Q_N := \{0, \dots, N-1\}^3 \quad (4)$$

is a discrete cube consisting of N^3 sites of \mathbb{Z}^3 . The particular location of the cube does not affect the limit in (3) in view of the homogeneity of the random field ξ .

For what follows, we assume that the particle system is constrained by the *nearest neighbour exclusion* (NNE) which forbids any two neighbouring sites of \mathbb{Z}^3 to be simultaneously occupied:

$$\begin{aligned} \mathbf{P}(A \neq 0, B \neq 0) &= \mathbf{P}(A \neq 0, C \neq 0) \\ &= \mathbf{P}(A \neq 0, E \neq 0) \\ &= 0. \end{aligned} \quad (5)$$

In view of the NNE constraint, the probability that two nearest sites of the lattice are both vacant is then computed by the inclusion-exclusion principle as

$$\begin{aligned} \mathbf{P}(A = B = 0) &= \mathbf{P}(A = 0) + \mathbf{P}(B = 0) \\ &\quad + \mathbf{P}(A \neq 0, B \neq 0) - 1 \\ &= 2S_0 - 1 \end{aligned} \quad (6)$$

(see also Eq. (4) of [1]), where S_0 is the vacancy probability in accordance with (2). Therefore, $S_0 \geq \frac{1}{2}$, with the extreme value $S_0 = \frac{1}{2}$ corresponding to the densest packing in \mathbb{Z}^3 , where one of any two nearest sites of the lattice is occupied while the other is vacant [22].

III. OCCUPIED SITE CONFIGURATIONS

Although the subsequent discussions are applicable to more general disordered condensed media, the interpretation of occupied site configurations in (1) will be aimed at a particular class of chemical systems. More precisely, we interpret each $w \in W$ as a spatial arrangement of a *coordination entity* which consists of a *central atom*, occupying a site of the simple cubic lattice, and an array of ligands bound to it. The ligands are not necessarily accommodated by \mathbb{Z}^3 and may reside

at sites of an interstitial lattice. However, their relative positions with respect to the central atom are specified for each occupied site configuration. Furthermore, the ligands are allowed to be shared by several “overlapping” coordination entities.

Let n denote the number of constituent particle species in the system. Within each of the species, particles are identical. For every $i = 1, \dots, n$, the i th species is endowed with a *coordination number* ν_i in the sense that any representative of the species is always shared by ν_i coordination entities associated with distinct sites of the simple cubic lattice.

We use the convention that the 1st species, which is further referred to as the *central species*, is represented only by central atoms and has coordination number $\nu_1 = 1$. More precisely, for each occupied site configuration $w \in W$, the corresponding coordination entity contains a single representative of the 1st species and the particle is the central atom in the entity.

For example, in the structural model of a binary silicate melt $\text{SiO}_2 - \text{M}_2\text{O}$ described in Section IV of [1], where M is a univalent metal, the elements of the set W represent 79 second nearest neighbour bonds (SNNBs) $X - O - Y$. This model involves $n = 3$ species, with oxygen being the central species. Considering that Si and M are the 2nd and 3rd particle species, their coordination numbers are $\nu_2 = 4$ and $\nu_3 = 1$, respectively.

In view of the spatial ergodicity hypothesis of Section II, the numbers of particles of different species in a macroscopically large fragment of the lattice, consisting of \mathcal{N} sites, are asymptotically given by

$$\mathcal{N}_i \sim \frac{\mathcal{N}}{\nu_i} \sum_{w \in W} \Gamma_{iw} S_w, \quad \text{as } \mathcal{N} \rightarrow +\infty, \quad i = 1, \dots, n. \quad (7)$$

Here, Γ_{iw} denotes the number of particles of the i th species in a site configuration w , and the denominator ν_i comes from the fact that particles of the i th species are counted ν_i times according to their coordination number. In particular, the number of particles of the central species in the lattice fragment is asymptotically given by

$$\mathcal{N}_1 \sim (1 - S_0)\mathcal{N}, \quad (8)$$

where the conventions $\nu_1 = 1$ and $\Gamma_{1w} = 1$ for all $w \in W$ are used in combination with the probability that a lattice site is occupied:

$$\mathbf{P}(A \neq 0) = 1 - S_0 = \sum_{w \in W} S_w. \quad (9)$$

The relative mole fractions of the constituent particle species with reference to the central species can therefore be defined by

$$y_i := \frac{\mathcal{N}_i}{\mathcal{N}_1}, \quad (10)$$

so that $y_1 = 1$. In what follows, the quantity y_i will be referred to as the *centralized* mole fraction of the i th particle

species. By dividing both parts of (7) by those of (8), it follows that (10) is equivalent to

$$\frac{1}{\nu_i} \sum_{w \in W} \Gamma_{iw} \tilde{S}_w = y_i, \quad i = 1, \dots, n. \quad (11)$$

Here, \tilde{S}_w denotes the conditional probability that a lattice site is at a state $w \in W$, provided this site is occupied. In accordance with (9) these conditional probabilities are computed as

$$\tilde{S}_w := \mathbf{P}(A = w \mid A \neq 0) = \frac{S_w}{1 - S_0}, \quad w \in W, \quad (12)$$

and comprise the conditional PMF $\tilde{S} := (\tilde{S}_w)_{w \in W} \in [0, 1]^W$ for the state of an occupied site of \mathbb{Z}^3 .

IV. EDGE-TO-SITE ENTROPY REDUCTION

In view of the NNE constraint (5), the common *edge PMF* $\mathbf{E} := (\mathbf{E}_{uv})_{u,v \in \Omega} \in [0, 1]^{\Omega^2}$ of the pairs (A, B) , (A, C) , (A, E) (which are regarded as Ω^2 -valued random variables) is given by

$$\begin{aligned} \mathbf{E}_{uv} &:= \mathbf{P}(A = u, B = v) \\ &= \mathbf{P}(A = u, C = v) \\ &= \mathbf{P}(A = u, E = v) \\ &= \begin{cases} 2S_0 - 1 & \text{if } u = v = 0 \\ S_u & \text{if } u \in W, v = 0 \\ S_v & \text{if } u = 0, v \in W \\ 0 & \text{if } u, v \in W \end{cases}, \end{aligned} \quad (13)$$

where use is also made of (6). As discussed in [1], the property that the edge PMF \mathbf{E} is specified completely by the site PMF \mathbf{S} is an important consequence of the NNE constraint [24] which allows the common *edge entropy*

$$\mathbf{H}_{\text{edge}} := - \sum_{u,v \in \Omega} \Lambda(\mathbf{E}_{uv}) \quad (14)$$

to be reduced to the site entropy

$$\mathbf{H}_{\text{site}} := - \sum_{u \in \Omega} \Lambda(S_u) \quad (15)$$

as

$$\mathbf{H}_{\text{edge}} = 2\mathbf{H}_{\text{site}} - \varphi(S_0), \quad (16)$$

see also Eq. (9) of [1]. Here, use is made of the shorthand notation

$$\Lambda(p) := p \ln p, \quad (17)$$

with the standard convention $\Lambda(0) := 0$, and the function φ is defined by

$$\varphi(p) := \Lambda(2p - 1) - 2\Lambda(p) = \Phi(p, p), \quad (18)$$

where

$$\Phi(p, q) := \Lambda(p + q - 1) - \Lambda(p) - \Lambda(q). \quad (19)$$

The graph of φ can be found in Fig. 2 of [1]. The quantity $\Phi(p, q)$ coincides with the Shannon mutual information [12, p. 19–22]

$$\mathbf{I}(\alpha; \beta) := \mathbf{H}(\alpha) + \mathbf{H}(\beta) - \mathbf{H}(\alpha, \beta)$$

between Boolean random variables α and β , which take values in the set $\{0, 1\}$ with probabilities $p := \mathbf{P}(\alpha = 0) = 1 - \mathbf{P}(\alpha = 1)$ and $q := \mathbf{P}(\beta = 0) = 1 - \mathbf{P}(\beta = 1)$ and are mutually exclusive in the sense that their Sheffer stroke $\alpha \bar{\wedge} \beta := \neg(\alpha \wedge \beta)$ (that is, the negated conjunction) takes the truth value almost surely: $\mathbf{P}(\alpha \bar{\wedge} \beta = 1) = 1$. The latter condition implies that $p + q \geq 1$.

Hence, by denoting the indicator function of the set W of occupied site configurations in (1) by $\mathcal{I}_W(\cdot)$, the quantity $\varphi(S_0)$ in (18) can be represented in terms of the Shannon information as

$$\begin{aligned} \varphi(S_0) &= \mathbf{I}(\mathcal{I}_W(A); \mathcal{I}_W(B)) \\ &= \mathbf{I}(\mathcal{I}_W(A); \mathcal{I}_W(C)) \\ &= \mathbf{I}(\mathcal{I}_W(A); \mathcal{I}_W(E)), \end{aligned}$$

where the Boolean random variables $\mathcal{I}_W(A), \dots, \mathcal{I}_W(H)$ are the *site occupancy indicators* (SOIs) introduced by Eq. (24) of [1].

Furthermore, the *edge-to-site entropy reduction*, described by (16), implies that $\varphi(S_0)$ is also the Shannon information

$$\begin{aligned} \varphi(S_0) &= 2\mathbf{H}_{\text{site}} - \mathbf{H}_{\text{edge}} \\ &= \mathbf{I}(A; B) = \mathbf{I}(A; C) = \mathbf{I}(A; E) \end{aligned}$$

between the states of neighbouring lattice sites which is induced by the NNE constraint. Indeed, if the states of such sites were independent random variables, the edge entropy would be equal to $2\mathbf{H}_{\text{site}}$.

V. FACE-TO-DIAGONAL ENTROPY REDUCTION

Similarly to the edge-to-site entropy reduction, NNE allows each of the *face entropies*

$$\mathbf{H}_{\text{face}}^{(12)} := \mathbf{H}(A, B, C, D), \quad (20)$$

$$\mathbf{H}_{\text{face}}^{(13)} := \mathbf{H}(A, B, E, F), \quad (21)$$

$$\mathbf{H}_{\text{face}}^{(23)} := \mathbf{H}(A, C, E, G), \quad (22)$$

to be reduced to the corresponding pair of *face-diagonal entropies*

$$\mathbf{H}_{\text{diag}}^{(12 \setminus)} := \mathbf{H}(A, D), \quad \mathbf{H}_{\text{diag}}^{(12 /)} := \mathbf{H}(B, C), \quad (23)$$

$$\mathbf{H}_{\text{diag}}^{(13 \setminus)} := \mathbf{H}(A, F), \quad \mathbf{H}_{\text{diag}}^{(13 /)} := \mathbf{H}(B, E), \quad (24)$$

$$\mathbf{H}_{\text{diag}}^{(23 \setminus)} := \mathbf{H}(A, G), \quad \mathbf{H}_{\text{diag}}^{(23 /)} := \mathbf{H}(C, E), \quad (25)$$

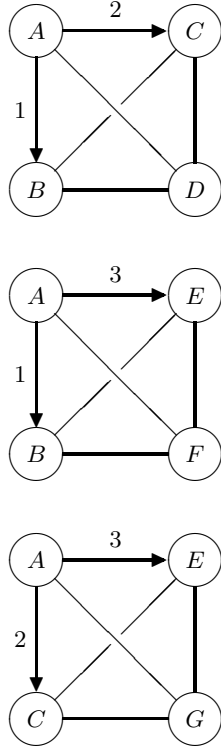


FIG. 2: The states (A, B, C, D) , (A, B, E, F) and (A, C, E, G) of three faces of the cubic cell in Fig. 1. The arrows labeled by 1, 2 and 3 specify the spanning coordinate axes for each of the faces. The face diagonals are depicted as thin straight lines and are mnemonically referenced by triples $(jk\ell)$ (with commas omitted for brevity), where $1 \leq j < k \leq 3$ and $\ell = \setminus, /$, and the end sites of the face diagonals are ordered “from left to right”. For example, the states of the face-diagonals $(13/)$ and $(23\setminus)$ are the pairs (B, E) and (A, G) , respectively.

where the meaning of the superscripts is elucidated by Fig. 2. Here, the face entropies are defined by

$$\mathbf{H}_{\text{face}}^{(jk)} := - \sum_{u,v,x,y \in \Omega} \Lambda(\mathbf{F}_{uvxy}^{(jk)}), \quad (26)$$

where the degree four arrays $\mathbf{F}^{(jk)} := (\mathbf{F}_{uvxy}^{(jk)})_{u,v,x,y \in \Omega}$ describe the joint PMFs of the quadruples (A, B, C, D) , (A, B, E, F) and (A, C, E, G) by

$$\mathbf{F}_{abcd}^{(12)} := \mathbf{P}(A = a, B = b, C = c, D = d), \quad (27)$$

$$\mathbf{F}_{abef}^{(13)} := \mathbf{P}(A = a, B = b, E = e, F = f), \quad (28)$$

$$\mathbf{F}_{aceg}^{(23)} := \mathbf{P}(A = a, C = c, E = e, G = g). \quad (29)$$

The array $\mathbf{F}^{(jk)}$ is referred to as the (j, k) th *face PMF*. Accordingly, the face-diagonal entropies in (23)–(25) are defined by

$$\mathbf{H}_{\text{diag}}^{(jk\ell)} := - \sum_{u,y \in \Omega} \Lambda(\mathbf{D}_{uy}^{(jk\ell)}), \quad (30)$$

where the matrices $\mathbf{D}^{(jk\ell)} = (\mathbf{D}_{uy}^{(jk\ell)})_{u,y \in \Omega}$ describe the corresponding *face-diagonal PMFs* by

$$\mathbf{D}_{ad}^{(12\setminus)} := \mathbf{P}(A = a, D = d), \quad (31)$$

$$\mathbf{D}_{bc}^{(12/)} := \mathbf{P}(B = b, C = c), \quad (32)$$

$$\mathbf{D}_{af}^{(13\setminus)} := \mathbf{P}(A = a, F = f), \quad (33)$$

$$\mathbf{D}_{be}^{(13/)} := \mathbf{P}(B = b, E = e), \quad (34)$$

$$\mathbf{D}_{ag}^{(23\setminus)} := \mathbf{P}(A = a, G = g), \quad (35)$$

$$\mathbf{D}_{ce}^{(23/)} := \mathbf{P}(C = c, E = e). \quad (36)$$

Associated with the (j, k) th face in Fig. 2 are two *face diagonal occupancy indicators* (FDOIs) $\omega^{(jk\setminus)}$ and $\omega^{(jk/)}$. Each of them is a Boolean random variable which indicates whether at least one of the end sites of the corresponding face diagonal is occupied. More precisely,

$$\omega^{(12\setminus)} := \mathcal{I}_W(A) \vee \mathcal{I}_W(D), \quad (37)$$

$$\omega^{(12/)} := \mathcal{I}_W(B) \vee \mathcal{I}_W(C), \quad (38)$$

$$\omega^{(13\setminus)} := \mathcal{I}_W(A) \vee \mathcal{I}_W(F), \quad (39)$$

$$\omega^{(13/)} := \mathcal{I}_W(B) \vee \mathcal{I}_W(E), \quad (40)$$

$$\omega^{(23\setminus)} := \mathcal{I}_W(A) \vee \mathcal{I}_W(G), \quad (41)$$

$$\omega^{(23/)} := \mathcal{I}_W(C) \vee \mathcal{I}_W(E), \quad (42)$$

where \vee denotes the logical disjunction, and use is made of the SOIs $\mathcal{I}_W(A), \dots, \mathcal{I}_W(H)$ mentioned in Section IV; see also Fig. 2.

The NNE constraint implies that the FDOIs, associated with any given face of the cubic cell, are also mutually exclusive. That is,

$$\omega^{(jk\setminus)} \bar{\wedge} \omega^{(jk/)} = 1 \quad (43)$$

for all $1 \leq j < k \leq 3$. Moreover, this property can be used as an equivalent reformulation of the NNE constraint. Indeed, the equivalence between the FDOI exclusion (43) and NNE follows from the identity $(\alpha \vee \delta) \bar{\wedge} (\beta \vee \gamma) = (\alpha \bar{\wedge} \beta) \wedge (\alpha \bar{\wedge} \gamma) \wedge (\beta \bar{\wedge} \delta) \wedge (\gamma \bar{\wedge} \delta)$ for Boolean variables $\alpha, \beta, \gamma, \delta$.

Therefore, by considering admissible *face configurations* which satisfy NNE, it can be shown that the face PMFs in (27)–(29) are expressed in terms of the face-diagonal PMFs in (31)–(36) as

$$\mathbf{F}_{0000}^{(jk)} = \mathbf{D}_{00}^{(jk\setminus)} + \mathbf{D}_{00}^{(jk/)} - 1, \quad (44)$$

$$\mathbf{F}_{u000}^{(jk)} = \mathbf{D}_{u0}^{(jk\setminus)}, \quad (45)$$

$$\mathbf{F}_{00y}^{(jk)} = \mathbf{D}_{0y}^{(jk\setminus)}, \quad (46)$$

$$\mathbf{F}_{u00y}^{(jk)} = \mathbf{D}_{uy}^{(jk\setminus)}, \quad (47)$$

$$\mathbf{F}_{0v00}^{(jk)} = \mathbf{D}_{v0}^{(jk/)}, \quad (48)$$

$$\mathbf{F}_{00x0}^{(jk)} = \mathbf{D}_{0x}^{(jk/)}, \quad (49)$$

$$\mathbf{F}_{0vx0}^{(jk)} = \mathbf{D}_{vx}^{(jk/)}, \quad u, v, x, y \in W. \quad (50)$$

Hence, in order to establish the above mentioned reduction of the face entropies to the face-diagonal entropies, it now remains to substitute (44)–(50) into the right-hand side of (26), so that

$$\begin{aligned} \mathbf{H}_{\text{face}}^{(jk)} &= \mathbf{H}_{\text{diag}}^{(jk\backslash)} + \Lambda(D_{00}^{(jk\backslash)}) \\ &\quad + \mathbf{H}_{\text{diag}}^{(jk/)} + \Lambda(D_{00}^{(jk/)}) \\ &\quad - \Lambda(F_{0000}^{(jk)}) \\ &= \mathbf{H}_{\text{diag}}^{(jk\backslash)} + \mathbf{H}_{\text{diag}}^{(jk/)} - \Phi^{(jk)}. \end{aligned} \quad (51)$$

Here, use is also made of (30) and (19) together with the Shannon information between the FDOIs associated with the (j, k) th face of the cubic cell:

$$\begin{aligned} \Phi^{(jk)} &:= \Phi(D_{00}^{(jk\backslash)}, D_{00}^{(jk/)}) \\ &= \mathbf{I}(\omega^{(jk\backslash)}; \omega^{(jk/)}). \end{aligned} \quad (52)$$

The *face-to-diagonal entropy reduction*, described by (51), implies that the quantity $\Phi^{(jk)}$ is the NNE-induced Shannon information not only between the FDOIs $\omega^{(jk\backslash)}$ and $\omega^{(jk/)}$ but also between the states of diagonals of the (j, k) th face in Fig. 2:

$$\Phi^{(12)} = \mathbf{I}(A, D; B, C), \quad (53)$$

$$\Phi^{(13)} = \mathbf{I}(A, F; B, E), \quad (54)$$

$$\Phi^{(23)} = \mathbf{I}(A, G; C, E). \quad (55)$$

VI. MARGINALIZATION

The homogeneity of the equilibrium random field ξ implies that, for every $1 \leq j < k \leq 3$, the (j, k) th face PMF $F^{(jk)}$, given by (27)–(29), marginalizes to the common edge PMF E in (13) as

$$F_{uv\oplus\oplus}^{(jk)} = E_{uv}, \quad (56)$$

$$F_{\oplus\oplus xy}^{(jk)} = E_{xy}, \quad (57)$$

$$F_{u\oplus x\oplus}^{(jk)} = E_{ux}, \quad (58)$$

$$F_{\oplus v\oplus y}^{(jk)} = E_{vy}, \quad u, v, x, y \in \Omega. \quad (59)$$

Here, for the sake of brevity, the subscript \oplus denotes the independent summation over the corresponding dimension of an array. For example, $F_{uv\oplus\oplus}^{(jk)} := \sum_{x,y \in \Omega} F_{uvxy}^{(jk)}$. The relations (56) and (57) imply that $F_{uv\oplus\oplus}^{(jk)} = F_{\oplus\oplus uv}^{(jk)}$, which reflects the local invariance of the face PMF under translations along coordinate axis k . Similarly, the equality $F_{u\oplus x\oplus}^{(jk)} = F_{\oplus u\oplus x}^{(jk)}$, which follows from (58) and (59), is related to the translational invariance along axis j .

By substituting (44)–(50) for the face PMF and (13) for the edge PMF, it follows that, in the NNE setting, (56)–(59) are equivalent to the marginalization of the face-diagonal PMFs to the site PMF

$$D_{u\oplus}^{(jk\ell)} = D_{\oplus u}^{(jk\ell)} = S_u, \quad u \in \Omega, \quad (60)$$

for all $1 \leq j < k \leq 3$ and $\ell = \backslash, /$, in combination with the inequality

$$D_{00}^{(jk\backslash)} + D_{00}^{(jk/)} \geq 1. \quad (61)$$

This inequality originates from (44) and couples the two (otherwise independent) sets of linear equations in (60) for different values of $\ell = \backslash, /$, according to which the row and column sums of the matrices $D^{(jk\backslash)}$ and $D^{(jk/)}$ reproduce the vector S .

VII. FACE-DIAGONAL ADMISSIBILITY

Some of the face-diagonal configurations may be prohibited as a result of structural restrictions, additional to NNE, such as spatial compatibility of coordination entities at the end sites of the face diagonal (for example, if the entities overlap). Irrespective of the nature of these additional geometric constraints, they are described by six Boolean matrices of *face-diagonal admissibility*

$$A^{(jk\ell)} := (A_{uv}^{(jk\ell)})_{u,v \in \Omega}, \quad 1 \leq j < k \leq 3, \ell = \backslash, /. \quad (62)$$

Here, $A_{uv}^{(jk\ell)}$ indicates whether the configuration (u, v) is allowed for the (j, k, ℓ) th face-diagonal; see Fig. 2. In particular, since the vacancy-vacancy configurations are admissible, then

$$A_{00}^{(jk\ell)} = 1.$$

Each of the face-diagonal PMFs $D^{(jk\ell)}$ is *dominated* by the corresponding admissibility matrix $A^{(jk\ell)}$ in (62) in the sense that their entries satisfy the implication

$$A_{uv}^{(jk\ell)} = 0 \implies D_{uv}^{(jk\ell)} = 0. \quad (63)$$

In probability theoretic terms, this means the absolute continuity [23] of $D^{(jk\ell)}$ with respect to $A^{(jk\ell)}$ which is also considered as a measure on the set Ω^2 . Equivalently, $D^{(jk\ell)}$ is concentrated on the set

$$\Omega^{(jk\ell)} := \{(u, v) \in \Omega^2 : A_{uv}^{(jk\ell)} = 1\} \quad (64)$$

which consists of admissible configurations for the (j, k, ℓ) th face diagonal.

VIII. FACE-TO-DIAGONAL REDUCTION OF KRAMERS-WANNIER APPROXIMATION

The entropy of the equilibrium random field ξ per site of the simple cubic lattice is defined by

$$\mathbf{h} := \lim_{N \rightarrow +\infty} \frac{\mathbf{H}(\xi_{Q_N})}{N^3}, \quad (65)$$

where $\xi_{Q_N} := (\xi_z)_{z \in Q_N}$ is the restriction of ξ to the discrete cube Q_N given by (4). Following the terminology of [11], we will refer to \mathbf{h} as the *entropy density* in order to emphasize

that the configurational entropy in (65) is averaged per lattice site, similarly to (3). An upper bound for \mathbf{h} is provided by the site entropy \mathbf{H}_{site} which corresponds to the Bragg-Williams approximation [4].

Recalling (14), (15) and (20)–(22), the Cluster Variation Method (CVM) with faces of cubic cells of \mathbb{Z}^3 as basic clusters and their edges and sites as subclusters, known as the Kramers-Wannier approximation [4], employs the entropy density estimate

$$\hat{\mathbf{h}}_{\text{face}} := 7\mathbf{H}_{\text{site}} - 9\mathbf{H}_{\text{edge}} + \sum_{1 \leq j < k \leq 3} \mathbf{H}_{\text{face}}^{(jk)}. \quad (66)$$

The latter takes into account the structure of the simple cubic lattice and the equilibrium statistical correlations in the particle system within faces of cubic cells through the appropriate Kikuchi-Barker coefficients [9] weighting the subcluster entropies. In contrast to \mathbf{H}_{site} , the quantity $\hat{\mathbf{h}}_{\text{face}}$ is not necessarily an upper bound for the entropy density in (65) nor is prevented from taking negative values. However, its advantage, even in comparison with the Bethe-Peierls entropy density estimate, is that $\hat{\mathbf{h}}_{\text{face}}$ “captures” more distant spatial correlations.

Due to the edge-to-site and face-to-diagonal entropy reductions in (16) and (51), which hold in the NNE setting, (66) takes the form

$$\begin{aligned} \hat{\mathbf{h}}_{\text{face}} = & 9\varphi(S_0) - 11\mathbf{H}_{\text{site}} \\ & + \sum_{1 \leq j < k \leq 3} \left(-\Phi^{(jk)} + \sum_{\ell = \backslash, /} \mathbf{H}_{\text{diag}}^{(jk\ell)} \right), \end{aligned} \quad (67)$$

which describes a *face-to-diagonal reduction* of the Kramers-Wannier entropy density approximation in the NNE framework, where use is also made of (52)–(55).

IX. INTERNAL ENERGY

The internal energy of the particle system is modelled by the sum of interaction energies over face diagonals of cubic cells. The energetics is parameterized by the six energy matrices

$$\mathbf{U}^{(jk\ell)} := (\mathbf{U}_{uv}^{(jk\ell)})_{u,v \in \Omega}, \quad 1 \leq j < k \leq 3, \ell = \backslash, /. \quad (68)$$

Here, $\mathbf{U}_{uv}^{(jk\ell)}$ denotes the interaction energy which is ascribed to the configuration (u, v) for the (j, k, ℓ) th face diagonal, with

$$\mathbf{U}_{00}^{(jk\ell)} := 0, \quad (69)$$

so that *vacancy-vacancy* configurations are endowed with zero energy. Hence, the average internal energy of the particle system per lattice site is computed as

$$\mathbf{U} := \sum_{1 \leq j < k \leq 3, \ell = \backslash, /} \langle \mathbf{U}^{(jk\ell)}, \mathbf{D}^{(jk\ell)} \rangle, \quad (70)$$

in terms of the six-tuple

$$\mathbf{D} := \left\{ \mathbf{D}^{(jk\ell)} : 1 \leq j < k \leq 3, \ell = \backslash, / \right\} \quad (71)$$

of the face-diagonal PMFs from (31)–(36), with $\langle \cdot, \cdot \rangle$ denoting the Frobenius inner product of matrices:

$$\langle \mathbf{U}^{(jk\ell)}, \mathbf{D}^{(jk\ell)} \rangle = \sum_{u,y \in \Omega} \mathbf{U}_{uy}^{(jk\ell)} \mathbf{D}_{uy}^{(jk\ell)}.$$

In particular, if the energy matrices in (68) are representable as

$$\mathbf{U}_{uy}^{(jk\ell)} = \frac{\mathbf{U}_u^{\text{site}} + \mathbf{U}_y^{\text{site}}}{12}, \quad (72)$$

where $\mathbf{U} := (\mathbf{U}_a)_{a \in \Omega}$ is a single site energy function satisfying $\mathbf{U}_0 = 0$, then (70) reduces to

$$\mathbf{U} = \langle \mathbf{U}, \mathbf{S} \rangle = \sum_{w \in W} \mathbf{U}_w^{\text{site}} \mathbf{S}_w. \quad (73)$$

The latter internal energy density corresponds to the model which is considered in [1]. Note that the denominator on the right-hand side of (72) takes into account the property that every site of the simple cubic lattice is shared by twelve face diagonals.

X. GIBBS ENERGY ESTIMATE

In the framework of the NNE-induced face-to-diagonal reduction of the Kramers-Wannier entropy density estimate described by (67) and with the internal energy density given by (70), the Gibbs free energy of the particle system per *occupied* site of the simple cubic lattice or, equivalently, per central atom, is approximated by

$$\hat{\mathbf{g}}_{\text{face}} := \min_{\mathbf{D}, \mathbf{S}} \frac{\mathbf{U} + P\mathbf{V}_1 - k_B T \hat{\mathbf{h}}_{\text{face}}}{1 - S_0}, \quad (74)$$

where the denominator originates from (8). Here, the minimum is taken over the six-tuple \mathbf{D} of face-diagonal PMFs in (71), satisfying the marginalization and admissibility constraints (60), (61) and (63), and over the site PMF \mathbf{S} subject to the balance equations (11) for given values of the pressure P , absolute temperature T and centralized mole fractions y_2, \dots, y_n in (10).

Furthermore, \mathbf{V}_1 in (74) denotes the physical volume of the cubic cell of the carrier lattice and is assumed to be constant, so that $\mathbf{U} + P\mathbf{V}_1$ is the *enthalpy density* of the particle system per lattice site. The Boltzmann constant $k_B = 1.381 \times 10^{-23} \text{ J/K}$ converts the information theoretic $\hat{\mathbf{h}}_{\text{face}}$ to the thermodynamic entropy density estimate $k_B \hat{\mathbf{h}}_{\text{face}}$, with only the *configurational* part of the entropy being taken into account.

The solution of the constrained minimization problem (74) will be described under an isotropy assumption of the next section, which further decreases the dimensionality of the problem by an order of magnitude.

XI. ISOTROPY EQUIVALENCE CLASSES

In addition to the NNE constraint, suppose the six-tuple of energy matrices in (68) is isotropic, that is, invariant with respect to the full octahedral symmetry group O_h of the cube [13] consisting of 48 isometries of the simple cubic lattice generated by mirror reflections and discrete rotations. In this case, the energy matrices $U^{(jkl)}$ are obtained by permuting the entries of $U^{(12\setminus)}$ and are parameterized by the interaction energies $\mathcal{U}_0, \dots, \mathcal{U}_d$ which are ascribed to elements of the isotropy equivalence classes

$$\Omega_0^{\text{diag}}, \dots, \Omega_d^{\text{diag}} \quad (75)$$

into which the set $\Omega^{(12\setminus)}$ of admissible configurations for the $(1, 2, \setminus)$ th face diagonal in (64) is split by the action of the isometry group O_h . More precisely,

$$U_{uv}^{(12\setminus)} = \mathcal{U}_i, \quad (u, v) \in \Omega_i^{\text{diag}}, \quad i = 0, \dots, d, \quad (76)$$

where

$$\Omega_0^{\text{diag}} := \{(0, 0)\} \quad (77)$$

is a singleton consisting of the vacancy-vacancy configuration, so that

$$\mathcal{U}_0 = 0$$

in accordance with (69). The equivalence classes $\Omega_1^{\text{diag}}, \dots, \Omega_d^{\text{diag}}$ are labelled in such a way that the first d_1 of them satisfy the inclusion

$$\bigcup_{i=1}^{d_1} \Omega_i^{\text{diag}} \subset (W \times \{0\}) \sqcup (\{0\} \times W) \quad (78)$$

(that is, they consist of face-diagonal configurations with precisely one occupied site), whereas the remaining $d_2 := d - d_1$ classes $\Omega_{d_1+1}^{\text{diag}}, \dots, \Omega_d^{\text{diag}}$ satisfy

$$\bigcup_{i=d_1+1}^d \Omega_i^{\text{diag}} \subset W^2 \quad (79)$$

and are formed from face-diagonal configurations with both sites occupied.

Assuming the absence of symmetry breaking, the isotropy of the energy matrices is inherited by the finite-dimensional probability distributions of the equilibrium random field ξ , including its site and face-diagonal PMFs. Under the isotropy assumption, the face-diagonal PMFs are permutations of entries of $D^{(12\setminus)}$ and are completely specified by the probabilities $\mathcal{D}_0, \dots, \mathcal{D}_d$ of representatives of the *face-diagonal classes* in (75) as

$$D_{uv}^{(12\setminus)} = \mathcal{D}_i, \quad (u, v) \in \Omega_i^{\text{diag}}, \quad i = 0, \dots, d. \quad (80)$$

In particular,

$$D_{00}^{(jkl)} = \mathcal{D}_0 \geq \frac{1}{2} \quad (81)$$

is the common probability of the vacancy-vacancy face-diagonal configurations, with the inequality being the isotropic version of (61).

By a similar reasoning, under the isotropy assumption, the site configurations are equiprobable within each of the equivalence classes

$$\Omega_0^{\text{site}}, \dots, \Omega_s^{\text{site}} \quad (82)$$

into which the site configurational space Ω in (1) is partitioned by the action of the isometry group O_h . More precisely, the site PMF S takes the form

$$S_u = \mathcal{S}_i, \quad u \in \Omega_i^{\text{site}}, \quad i = 0, \dots, s, \quad (83)$$

where \mathcal{S}_i is its common value at elements of the i th *site class* Ω_i^{site} . Here, $\Omega_0^{\text{site}} = \{0\}$ consists of the vacancy, so that $\mathcal{S}_0 = S_0$, while the remaining site classes partition the set of occupied site configurations:

$$\bigcup_{i=1}^s \Omega_i^{\text{site}} = W.$$

Therefore, in the isotropic case being considered, the internal energy density (70) and the entropy density estimate (67) take the form

$$\begin{aligned} U &= 6 \langle U^{(12\setminus)}, D^{(12\setminus)} \rangle \\ &= 6 \sum_{i=1}^d \delta_i \mathcal{U}_i \mathcal{D}_i, \end{aligned} \quad (84)$$

$$\begin{aligned} \hat{\mathbf{h}}_{\text{face}} &= 9\varphi(\mathcal{S}_0) - 11\mathbf{H}_{\text{site}} \\ &\quad - 3\varphi(\mathcal{D}_0) + 6\mathbf{H}_{\text{diag}}. \end{aligned} \quad (85)$$

Here, use has been made of (18), (19), (52), (76) and (81). Furthermore,

$$\mathbf{H}_{\text{diag}} := \mathbf{H}_{\text{diag}}^{(jkl)} = - \sum_{i=0}^d \delta_i \Lambda(\mathcal{D}_i) \quad (86)$$

is the common value of the face-diagonal entropies in (30), and

$$\mathbf{H}_{\text{site}} = - \sum_{i=0}^s \sigma_i \Lambda(\mathcal{S}_i), \quad (87)$$

where

$$\sigma_i := \#\Omega_i^{\text{site}}, \quad \delta_j := \#\Omega_j^{\text{diag}} \quad (88)$$

denote the cardinalities of the site and face-diagonal equivalence classes in (82) and (75), with $\sigma_0 = \delta_0 = 1$.

XII. CONDITIONED REPRESENTATION

In what follows, we will employ the decomposition [12, p. 44] of the joint entropy for discrete random variables η and

$\zeta := f(\eta)$, the second of which is a deterministic function of η :

$$\begin{aligned}\mathbf{H}(\eta) &= \mathbf{H}(\eta, \zeta) \\ &= \mathbf{H}(\zeta) + \mathbf{H}(\eta | \zeta) \\ &= \mathbf{H}(\zeta) + \sum_z \mathbf{P}(\zeta = z) \mathbf{H}(\eta | \zeta = z),\end{aligned}\quad (89)$$

where $\mathbf{H}(\cdot | \cdot)$ denotes the conditional entropy, and the sum is taken over the range of ζ . Similarly to the lines of reasoning for Eq. (25) in [1], the site entropy in (87) can be represented as

$$\mathbf{H}_{\text{site}} = \mathbf{H}_{\text{SOI}} + (1 - \mathcal{S}_0) \mathbf{H}(A | A \neq 0), \quad (90)$$

which is obtained by applying (89) to the random variable $\eta := A$ and the SOI $\zeta := \mathcal{I}_W(A)$ (see Section IV) and using the relation

$$\mathbf{H}(A | \mathcal{I}_W(A) = 0) = \mathbf{H}(A | A = 0) = 0.$$

In the framework of the isotropy assumption of Section XI, the common entropy of the SOIs $\mathcal{I}_W(A), \dots, \mathcal{I}_W(H)$ in (90) is computed as

$$\mathbf{H}_{\text{SOI}} := -\Lambda(\mathcal{S}_0) - \Lambda(1 - \mathcal{S}_0). \quad (91)$$

Also, the conditional entropy of the state of a given site of the carrier lattice, provided that the site is occupied, takes the form

$$\mathbf{H}(A | A \neq 0) = -\sum_{i=1}^s \sigma_i \Lambda(\tilde{\mathcal{S}}_i), \quad (92)$$

with

$$\tilde{\mathcal{S}}_i := \frac{\mathcal{S}_i}{1 - \mathcal{S}_0}, \quad i = 1, \dots, s, \quad (93)$$

describing the isotropic version of the conditional site PMF in (12) in view of (83) and (88), so that $\sum_{i=1}^s \sigma_i \tilde{\mathcal{S}}_i = 1$.

A similar conditioned representation for the face-diagonal entropy in (86) is

$$\begin{aligned}\mathbf{H}_{\text{diag}} &= \mathbf{H}_{\text{FDOI}} \\ &+ (1 - \mathcal{D}_0) \mathbf{H}(A, D | \omega^{(12\setminus)} = 1),\end{aligned}\quad (94)$$

where

$$\begin{aligned}\mathbf{H}_{\text{FDOI}} &:= \mathbf{H}(\omega^{(jk\ell)}) \\ &= -\Lambda(\mathcal{D}_0) - \Lambda(1 - \mathcal{D}_0)\end{aligned}\quad (95)$$

is the common entropy of the FDOIs in (37)–(42) under the isotropy assumption, and

$$\mathbf{H}(A, D | \omega^{(12\setminus)} = 1) = -\sum_{i=1}^d \delta_i \Lambda(\tilde{\mathcal{D}}_i) \quad (96)$$

is the conditional entropy of the state of a given face diagonal provided that the associated FDOI is true. Here, in accordance with (80), the probabilities

$$\tilde{\mathcal{D}}_i := \frac{\mathcal{D}_i}{1 - \mathcal{D}_0}, \quad i = 1, \dots, d, \quad (97)$$

comprise the conditional face-diagonal PMF in the isotropic case, so that $\sum_{i=1}^d \delta_i \tilde{\mathcal{D}}_i = 1$ in view of (88). Furthermore, the internal energy density in (84) is expressed in terms of $\tilde{\mathcal{D}}_1, \dots, \tilde{\mathcal{D}}_d$ as

$$\mathbf{U} = 6(1 - \mathcal{D}_0) \sum_{i=1}^d \delta_i \mathcal{U}_i \tilde{\mathcal{D}}_i. \quad (98)$$

XIII. BALANCE EQUATIONS

For any given particle species, the number of its representatives in a coordination entity depends only on the site class to which this entity belongs, so that

$$\Gamma_{iw} = \gamma_{ij}, \quad w \in \Omega_j^{\text{site}}.$$

Hence, under the isotropy assumption of Section XI, the balance equations (11) take the form

$$\frac{1}{\nu_i} \sum_{j=1}^s \gamma_{ij} \sigma_j \tilde{\mathcal{S}}_j = y_i, \quad i = 1, \dots, n, \quad (99)$$

where use is made of the conditional site PMF described by (93). By assembling the coefficients in (99) into an $(n \times s)$ -matrix

$$\Upsilon := \left(\frac{\gamma_{ij} \sigma_j}{\nu_i} \right)_{1 \leq i \leq n, 1 \leq j \leq s}, \quad (100)$$

this system of linear equations can be represented in vector-matrix form as

$$\Upsilon \tilde{\mathcal{S}} = Y. \quad (101)$$

Here, $\tilde{\mathcal{S}} := (\tilde{\mathcal{S}}_i)_{1 \leq i \leq s}$ and $Y := (y_i)_{1 \leq i \leq n}$ are column-vectors of the probabilities from (93) and the centralized mole fractions defined by (10). The first row of the matrix Υ consists of the occupied site class cardinalities:

$$\Upsilon_{1\bullet} = [\sigma_1 \ \dots \ \sigma_s].$$

XIV. MARGINALIZATION EQUATIONS

Under the isotropy assumption of Section XI, the marginalization equations (60) can be represented in vector-matrix form as

$$M\mathcal{D} = L\mathcal{S}, \quad (102)$$

where $\mathcal{D} := (\mathcal{D}_i)_{0 \leq i \leq d}$ and $\mathcal{S} := (\mathcal{S}_i)_{0 \leq i \leq s}$ are column-vectors with entries from (80) and (83). The matrix M is given by

$$M := \left[\begin{array}{c|c} 1 & \widetilde{M} \end{array} \right], \quad \widetilde{M} := \begin{bmatrix} M_0 \\ M_1 \end{bmatrix} \quad (103)$$

and is assumed to be of full row rank. Here, $\mathbf{0}_{p \times q}$ denotes the $(p \times q)$ -matrix of zeros. Also, the matrix L in (102) is block-diagonal,

$$L := \left[\begin{array}{cc} 1 & \mathbf{0}_{1 \times s} \\ \mathbf{0}_{r \times 1} & \widetilde{L} \end{array} \right], \quad \widetilde{L} := \begin{bmatrix} \mathbf{1}_{r_1} & & \mathbf{0} \\ & \ddots & \\ \mathbf{0} & & \mathbf{1}_{r_s} \end{bmatrix}, \quad (104)$$

where $\mathbf{1}_p$ denotes the p -dimensional column-vector of ones, and

$$r := \sum_{i=1}^s r_i.$$

The block M_0 of the matrix \widetilde{M} in (103) is a d -dimensional row-vector whose first d_1 entries are the half-cardinalities of the face-diagonal classes $\Omega_1^{\text{diag}}, \dots, \Omega_{d_1}^{\text{diag}}$ from (78) padded with d_2 zeros:

$$M_0 := \frac{1}{2} [\delta_1 \ \dots \ \delta_{d_1} \ \mathbf{0}_{1 \times d_2}]. \quad (105)$$

Accordingly, the bottom block M_1 is an $(r \times d)$ -matrix. The structure of the matrices M and L allows (102) to be represented in the form

$$\widetilde{M} \widetilde{\mathcal{D}} = \begin{bmatrix} 1 - \theta \\ \theta \widetilde{L} \widetilde{\mathcal{S}} \end{bmatrix} \quad (106)$$

whose right-hand side depends on the probabilities \mathcal{D}_0 and \mathcal{S}_0 only through the auxiliary variable

$$\theta := \frac{1 - \mathcal{S}_0}{1 - \mathcal{D}_0}. \quad (107)$$

Here, $\widetilde{\mathcal{D}} := (\widetilde{\mathcal{D}}_i)_{1 \leq i \leq d}$ is the column-vector of probabilities from (97). In view of (37), the probabilistic meaning of θ is clarified by

$$\begin{aligned} \theta &= \frac{\mathbf{P}(A \neq 0)}{\mathbf{P}(\omega^{(12 \setminus)} = 1)} \\ &= \frac{\mathbf{P}(A \neq 0, \omega^{(12 \setminus)} = 1)}{\mathbf{P}(\omega^{(12 \setminus)} = 1)} \\ &= \mathbf{P}(A \neq 0 \mid \omega^{(12 \setminus)} = 1), \end{aligned}$$

where use is made of the property that the event $A \neq 0$ implies $\omega^{(12 \setminus)} = 1$. In contrast to the pair $(\mathcal{D}_0, \mathcal{S}_0)$, whose admissible values are depicted in Fig. 3, the quantities \mathcal{D}_0 and θ are functionally independent, with each of them taking values in the interval $[1/2, 1)$.

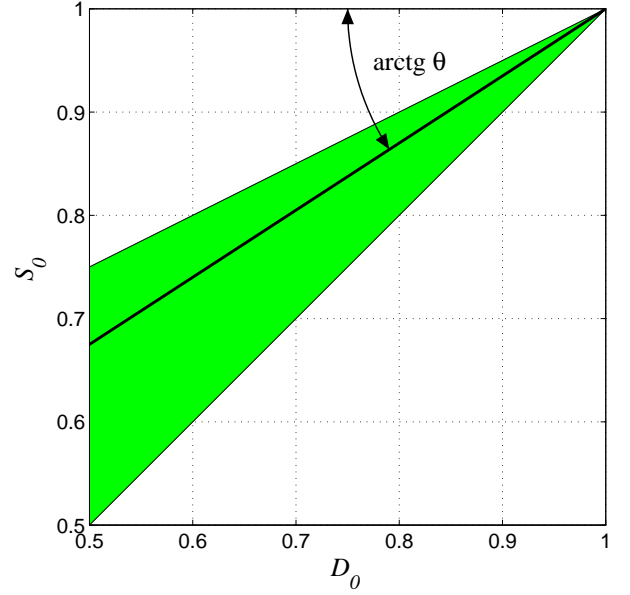


FIG. 3: The shaded triangle is the set of admissible values of the pair $(\mathcal{D}_0, \mathcal{S}_0)$, which is described by $\frac{1}{2} \leq \mathcal{D}_0 < \mathcal{S}_0 \leq \frac{1}{2}(1 + \mathcal{D}_0)$. The bold line segment represents the set of constant ratio θ defined by (107).

XV. SEPARATION OF VARIABLES

By dividing the numerator in (74) by $k_B T$, the minimization can be reduced to that of the following dimensionless function

$$\mathbf{g} := \frac{1}{1 - \mathcal{S}_0} \left(K + \frac{\mathbf{U}}{k_B T} - \widehat{\mathbf{h}}_{\text{face}} \right), \quad (108)$$

where

$$K := \frac{PV_1}{k_B T} \quad (109)$$

resembles the compressibility factor. In view of (107), the probability \mathcal{S}_0 can be expressed in terms of \mathcal{D}_0 and θ as

$$\mathcal{S}_0 = 1 - (1 - \mathcal{D}_0)\theta. \quad (110)$$

This will allow \mathbf{g} in (108) to be minimized as a function of $\mathcal{D}_0, \theta, \widetilde{\mathcal{D}}, \widetilde{\mathcal{S}}$ by employing a separation-of-variables technique, similar to that in Dynamic Programming. More precisely, \mathbf{g} can be split into the sum of two functions which share θ as a common argument:

$$\mathbf{g} = \mathbf{g}_0(\mathcal{D}_0, \theta) + \mathbf{g}_1(\theta, \widetilde{\mathcal{D}}, \widetilde{\mathcal{S}}). \quad (111)$$

Since \mathcal{D}_0 enters the balance and marginalization equations (101) and (106) only through θ , then the variables \mathcal{D}_0 and $(\widetilde{\mathcal{D}}, \widetilde{\mathcal{S}})$ are functionally independent for any given value of θ . Therefore, the problem of constrained minimization of \mathbf{g} in

(111) can be decomposed as

$$\begin{aligned}\widehat{\mathbf{g}} &:= \min_{\mathcal{D}_0, \theta, \tilde{\mathcal{D}}, \tilde{\mathcal{S}}} \mathbf{g} \\ &= \min_{1/2 \leq \theta < 1} (\widehat{\mathbf{g}}_0(\theta) + \widehat{\mathbf{g}}_1(\theta))\end{aligned}\quad (112)$$

into the optimization problems

$$\widehat{\mathbf{g}}_0(\theta) := \min_{1/2 \leq \mathcal{D}_0 < 1} \mathbf{g}_0(\mathcal{D}_0, \theta), \quad (113)$$

$$\begin{aligned}\widehat{\mathbf{g}}_1(\theta) &:= \min\{\mathbf{g}_1(\theta, \tilde{\mathcal{D}}, \tilde{\mathcal{S}}) : \\ &\quad \tilde{\mathcal{D}}, \tilde{\mathcal{S}} \text{ satisfy (101) and (106)}\}\end{aligned}\quad (114)$$

which are solved separately for any given value of θ as a “master” parameter. The specific form of the functions \mathbf{g}_0 and \mathbf{g}_1 is as follows. A combination of (85)–(88), (98) and (18) with (108) yields

$$\mathbf{g}_0(\mathcal{D}_0, \theta) = \frac{K + \psi(\mathcal{D}_0, 1 - (1 - \mathcal{D}_0)\theta)}{(1 - \mathcal{D}_0)\theta}, \quad (115)$$

where

$$\begin{aligned}\psi(\mathcal{D}_0, \mathcal{S}_0) &:= 3\varphi(\mathcal{D}_0) - 6\mathbf{H}_{\text{FDOI}} \\ &\quad - 9\varphi(\mathcal{S}_0) + 11\mathbf{H}_{\text{SOI}} \\ &= 3\Lambda(2\mathcal{D}_0 - 1) + 6\Lambda(1 - \mathcal{D}_0) \\ &\quad - 9\Lambda(2\mathcal{S}_0 - 1) + 7\Lambda(\mathcal{S}_0) \\ &\quad - 11\Lambda(1 - \mathcal{S}_0),\end{aligned}\quad (116)$$

and \mathcal{S}_0 is expressed in terms of \mathcal{D}_0 and θ by (110). A similar reasoning leads to

$$\begin{aligned}\mathbf{g}_1(\theta, \tilde{\mathcal{D}}, \tilde{\mathcal{S}}) &= \frac{6}{\theta} \sum_{i=1}^d \delta_i \left(\frac{\mathcal{U}_i \tilde{\mathcal{D}}_i}{k_B T} + \Lambda(\tilde{\mathcal{D}}_i) \right) \\ &\quad - 11 \sum_{j=1}^s \sigma_j \Lambda(\tilde{\mathcal{S}}_j).\end{aligned}\quad (117)$$

The pressure P in (109) and the central mole fractions in (10) enter the Gibbs free energy approximation in two separate ways — via the function \mathbf{g}_0 in (115) and through the linear constraints in (114) which specify the function $\widehat{\mathbf{g}}_1$. However, being “blended” by the minimization over θ on the right-hand side of (112), the effects of pressure and chemical composition are not additive.

A. Minimization of \mathbf{g}_0

A combination of (110) with (115) and (116) yields the following partial derivative of the function \mathbf{g}_0 with respect to the probability \mathcal{D}_0 :

$$\begin{aligned}\partial_{\mathcal{D}_0} \mathbf{g}_0 &= \frac{(1 - \mathcal{D}_0)(\partial_{\mathcal{D}_0} \psi + \theta \partial_{\mathcal{S}_0} \psi) + K + \psi}{(1 - \mathcal{D}_0)^2 \theta} \\ &= \frac{(1 - \mathcal{D}_0) \partial_{\mathcal{D}_0} \psi + (1 - \mathcal{S}_0) \partial_{\mathcal{S}_0} \psi + K + \psi}{(1 - \mathcal{D}_0)^2 \theta} \\ &= \frac{K + 3 \ln(2\mathcal{D}_0 - 1) - 9 \ln(2\mathcal{S}_0 - 1) + 7 \ln \mathcal{S}_0}{(1 - \mathcal{D}_0)^2 \theta}.\end{aligned}\quad (118)$$

For any given $\theta \in [1/2, 1)$, the numerator of the fraction in (118) tends to $-\infty$ as $\mathcal{D}_0 \rightarrow \frac{1}{2}$ and approaches the quantity $K > 0$ as $\mathcal{D}_0 \rightarrow 1$. Hence, by the Intermediate Value Theorem, this ensures solvability of the equation $\partial_{\mathcal{D}_0} \mathbf{g}_0 = 0$ on the interval and achievability of the minimum in (113). Moreover, the minimum is achieved at a unique point which is related to the appropriate root $\chi := 1 - \mathcal{D}_0 \in (0, 1/2]$ of the decic equation

$$\left(\chi - \frac{1}{2}\right)^3 \left(\chi - \frac{1}{\theta}\right)^7 + 64\theta^2 e^{-K} \left(\chi - \frac{1}{2\theta}\right)^9 = 0, \quad (119)$$

whose unique solvability is established in Appendix A. The corresponding values of the probabilities \mathcal{D}_0 and \mathcal{S}_0 are given by

$$\widehat{\mathcal{D}}_0 := 1 - \chi, \quad \widehat{\mathcal{S}}_0 := 1 - \chi\theta \quad (120)$$

and are depicted, together with $\widehat{\mathbf{g}}_0$, as functions of θ and K in Figs. 4 and 5. For any value of the parameter $K > 0$ in

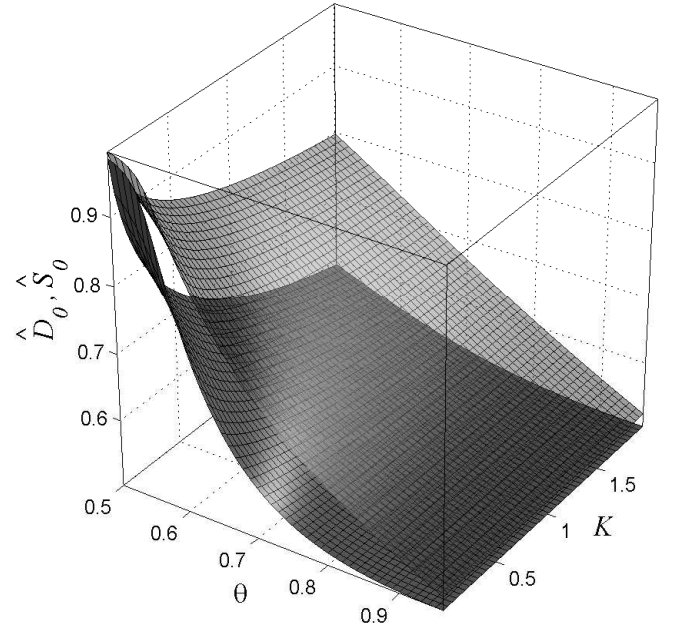


FIG. 4: The graphs of $\widehat{\mathcal{D}}_0$ (lower opaque surface) and $\widehat{\mathcal{S}}_0$ (upper transparent surface) as functions of θ and K defined by (119) and (120).

(109), the root χ is a smooth function of $\theta \in [1/2, 1)$ and so also are $\widehat{\mathcal{D}}_0$, $\widehat{\mathcal{S}}_0$ and $\widehat{\mathbf{g}}_0$. Since $\partial_{\mathcal{D}_0} \mathbf{g}_0$ vanishes at $\mathcal{D}_0 = \widehat{\mathcal{D}}_0$, then

$$\begin{aligned}\widehat{\mathbf{g}}'_0(\theta) &= \partial_{\theta} \mathbf{g}_0(\mathcal{D}_0, \theta) \Big|_{\mathcal{D}_0 = \widehat{\mathcal{D}}_0} \\ &= - \frac{(1 - \widehat{\mathcal{S}}_0) \partial_{\mathcal{S}_0} \psi + K + \psi}{(1 - \widehat{\mathcal{D}}_0) \theta^2}.\end{aligned}\quad (121)$$

From data on the molar volume of magmatic liquids in Fig. 2.3 on p. 29 of [16], a typical value of the cell volume is $V_1 \sim$

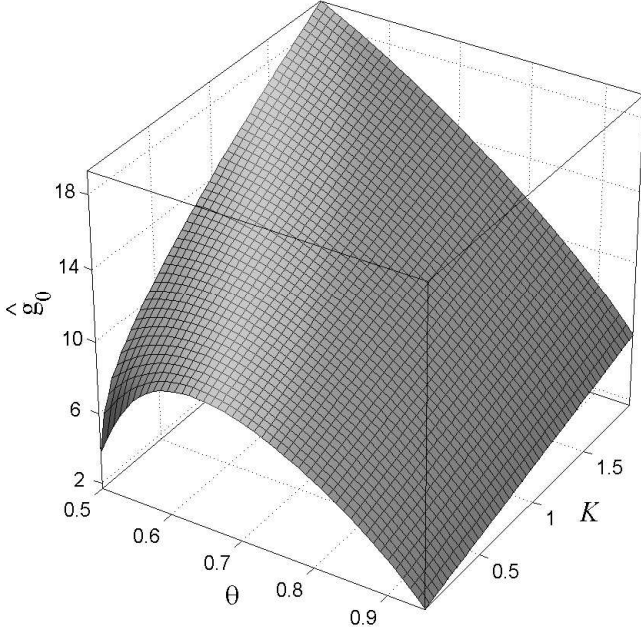


FIG. 5: The graph of \hat{g}_0 defined by (113) as a function of θ and K .

10^{-29}m^3 , and hence, an estimate of the quantity K in (109) at $P \sim 10^5\text{Pa}$ and $T \sim 10^3\text{K}$ is $K \sim 10^{-4}$. Thus, typically being small, K becomes close to 1 for tectonic pressures of 10^8 to 10^9 Pa.

B. Minimization of g_1

Although the function $g_1(\theta, \tilde{\mathcal{D}}, \tilde{\mathcal{S}})$ in (117) is strictly convex with respect to $\tilde{\mathcal{D}}$, it is strictly concave with respect to $\tilde{\mathcal{S}}$. This saddle-like landscape (where there is a guarantee only for a unique minimum over $\tilde{\mathcal{D}}$ at a given $\tilde{\mathcal{S}}$) complicates the constrained minimization problem in (114). The associated Lagrange function is given by

$$\begin{aligned} \mathcal{L}(\theta, \tilde{\mathcal{D}}, \tilde{\mathcal{S}}, \lambda_0, \lambda, \mu) = & g_1(\theta, \tilde{\mathcal{D}}, \tilde{\mathcal{S}}) \\ & - \lambda_0(M_0\tilde{\mathcal{D}} - (1 - \theta)) \\ & - \langle \lambda, M_1\tilde{\mathcal{D}} - \theta\tilde{L}\tilde{\mathcal{S}} \rangle \\ & - \langle \mu, \Upsilon\tilde{\mathcal{S}} - Y \rangle. \end{aligned} \quad (122)$$

Here, the scalar λ_0 and the column-vectors $\lambda := (\lambda_j)_{1 \leq j \leq r}$ and $\mu := (\mu_k)_{1 \leq k \leq n}$ are the Lagrange multipliers associated with the marginalization and balance equations (106) and (101), and use is made of the structure of the matrix \tilde{M} in (103). By substituting (117) into (122), it follows that the condition of stationarity of the Lagrange function with respect to

$\tilde{\mathcal{D}}$ takes the form

$$\begin{aligned} \partial_{\tilde{\mathcal{D}}_i} \mathcal{L} = & \frac{6\delta_i}{\theta} \left(\frac{\mathcal{U}_i}{k_B T} + 1 + \ln \tilde{\mathcal{D}}_i \right) \\ & - \lambda_0(M_0)_i - \langle \lambda, (M_1)_{\bullet i} \rangle = 0, \quad i = 1, \dots, d, \end{aligned} \quad (123)$$

where $(M_0)_i$ denotes the i th entry of the row-vector M_0 in (105) and $(M_1)_{\bullet i}$ is the i th column of the matrix M_1 . In a similar vein, the equations of stationarity of \mathcal{L} with respect to $\tilde{\mathcal{S}}$ are

$$\begin{aligned} \partial_{\tilde{\mathcal{S}}_j} \mathcal{L} = & -11\sigma_j \left(1 + \ln \tilde{\mathcal{S}}_j \right) \\ & + \theta \langle \lambda, \tilde{L}_{\bullet j} \rangle - \langle \mu, \Upsilon_{\bullet j} \rangle = 0, \quad j = 1, \dots, s, \end{aligned} \quad (124)$$

where $\Upsilon_{\bullet j}$ and $\tilde{L}_{\bullet j}$ denote the j th columns of the matrices Υ and \tilde{L} in (100) and (104), respectively. Hence, $\tilde{\mathcal{D}}$ and $\tilde{\mathcal{S}}$ are expressed in terms of the Lagrange multipliers λ_0 , λ and μ as

$$\begin{aligned} \tilde{\mathcal{D}}_i = & \exp \left(\frac{\theta}{6\delta_i} (\lambda_0(M_0)_i + \langle \lambda, (M_1)_{\bullet i} \rangle) - \frac{\mathcal{U}_i}{k_B T} - 1 \right), \\ \tilde{\mathcal{S}}_j = & \exp \left(\frac{1}{11\sigma_j} (\theta \langle \lambda, \tilde{L}_{\bullet j} \rangle - \langle \mu, \Upsilon_{\bullet j} \rangle) - 1 \right). \end{aligned}$$

Substitution of these expressions to (106) and (101) leads to a system of $1 + r + n$ nonlinear equations for the scalar variables $\lambda_0, \dots, \lambda_r$ and μ_1, \dots, μ_n , which can be solved numerically, for example, by Newton iterations (see Appendix B) organized as two nested loops. The inner loop iterates for λ_0 and λ in order to achieve (106), while the outer loop solves for μ to satisfy (101).

Note, however, that the development of a reliable algorithm for solving the constrained optimization problem in (114) (in particular, able to avoid false extrema) is a separate problem which needs to be solved for the computer implementation of this approach.

The minimum value in (114), delivered by a stationary point $(\hat{\lambda}_0, \hat{\lambda}, \hat{\mu})$, can be computed by multiplying both sides of (123) and (124) by $\tilde{\mathcal{D}}_i$ and $\tilde{\mathcal{S}}_j$, respectively, and taking the sum over i and j :

$$\hat{g}_1(\theta) = (1 - \theta)\hat{\lambda}_0 + \langle Y, \hat{\mu} \rangle + 11 - \frac{6}{\theta},$$

where use is also made of (117) and the previously mentioned identities $\sum_{i=1}^d \delta_i \tilde{\mathcal{D}}_i = \sum_{j=1}^s \sigma_j \tilde{\mathcal{S}}_j = 1$.

C. Minimization over the master parameter

If, in combination with the stationarity condition of the previous section, the minimum on the right-hand side of (112) is achieved at an interior value of the master parameter θ , then

$$\begin{aligned} \hat{g}'_0(\theta) + \partial_{\theta} \mathcal{L} = & \hat{g}'_0(\theta) + \partial_{\theta} g_1 \\ & - \hat{\lambda}_0 + \langle \hat{\lambda}, \hat{\tilde{L}}\hat{\tilde{\mathcal{S}}} \rangle = 0. \end{aligned} \quad (125)$$

TABLE I: The site classes $\Omega_0^{\text{site}}, \dots, \Omega_6^{\text{site}}$ for the NNE-constrained simple cubic lattice model of the $\text{SiO}_2\text{--M}_2\text{O}$ liquid silicate, with bond angles rounded to whole degrees. Also shown are the cardinalities of the classes and the O, Si and M contents in their representatives.

k	Ω_k^{site}	σ_k	γ_{1k}	γ_{2k}	γ_{3k}
0	oxygen vacancy	1	0	0	0
1	$\text{Si--O--Si} \angle 109^\circ$	12	1	2	0
2	$\text{Si--O--Si} \angle 180^\circ$	4	1	2	0
3	$\text{Si--O--M} \angle 55^\circ$	24	1	1	1
4	$\text{Si--O--M} \angle 125^\circ$	24	1	1	1
5	$\text{M--O--M} \angle 90^\circ$	12	1	0	2
6	$\text{M--O--M} \angle 180^\circ$	3	1	0	2

Here, $\hat{\mathbf{g}}'_0$ is given by (121), and the partial derivatives of the Lagrange function \mathcal{L} from (122) and those of the function \mathbf{g}_1 from (117) with respect to θ are evaluated at the solution $(\hat{\mathcal{D}}, \hat{\mathcal{S}})$ of the constrained minimization problem of the previous section and the associated Lagrange multipliers $(\hat{\lambda}_0, \hat{\lambda})$. Therefore, the above mentioned two-loop scheme can be equipped with an outermost loop which solves for the master parameter θ to satisfy (125).

XVI. AN APPLICATION TO MODELLING OF $\text{SiO}_2 - \text{M}_2\text{O}$

The isotropic version of the face-to-diagonal reduction of the Kramers-Wannier approximation is applicable to the statistical mechanical modeling of the $\text{SiO}_2\text{--M}_2\text{O}$ silicate melt in the framework of the above mentioned NNE cubic lattice structural model [1], where M is a univalent metal. The data below provide an insight into the combinatorial and numerical aspects of this application.

Including the oxygen vacancy, the set Ω consists of 80 site configurations, which are split into seven isotropy equivalence classes described in Table I. By a computer-aided analysis of configurational spaces on larger clusters (carried out using MATLAB), each of the sets $\Omega^{(jkl)}$ is split into 384 isotropy equivalence classes and consists of 2711 admissible face-diagonal configurations, which is significantly less than $(\#\Omega)^2 = 6400$ due to the geometric constraints described in Section IV of [1]; see also Fig. 6.

Furthermore, 118 admissible *oxygen-vacancy* face-diagonal configurations, which belong to the union of the sets on the right-hand side of (78), are split into $d_1 = 20$ isotropy equivalence classes $\Omega_1^{\text{diag}}, \dots, \Omega_{20}^{\text{diag}}$. The remaining 2592 *oxygen-oxygen* face-diagonal configurations belonging to W^2 are split into $d_2 = 363$ equivalence classes $\Omega_{21}^{\text{diag}}, \dots, \Omega_{383}^{\text{diag}}$. Among them, 184 face-diagonal configurations, partitioned into 31 equivalence class, are *overlapping* coordination entities which share a common Si atom and, thus, can be written as X--O--Si--O--Y , where X and Y stand for Si or M cations; see Figs. 7 and 8.

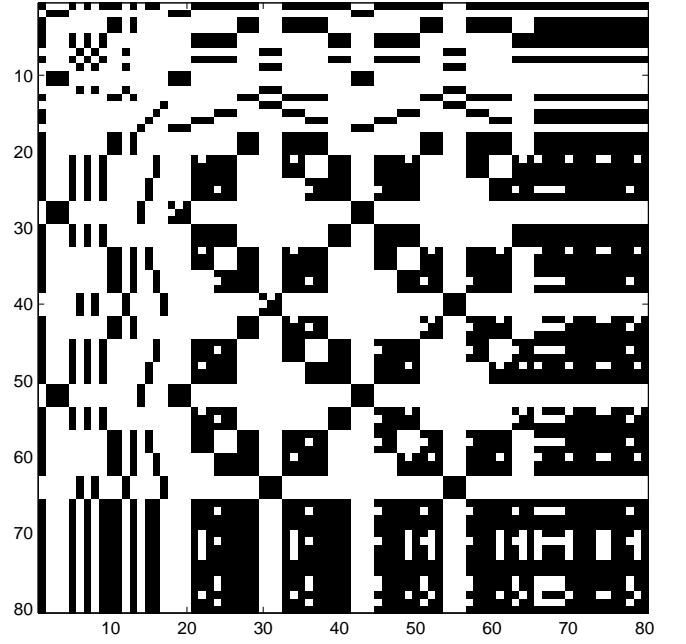


FIG. 6: The sparsity pattern of the face-diagonal admissibility matrix $\mathbf{A}^{(12\setminus)}$ for the $\text{SiO}_2\text{--M}_2\text{O}$ silicate model, with its 2711 nonzero entries shown in black.

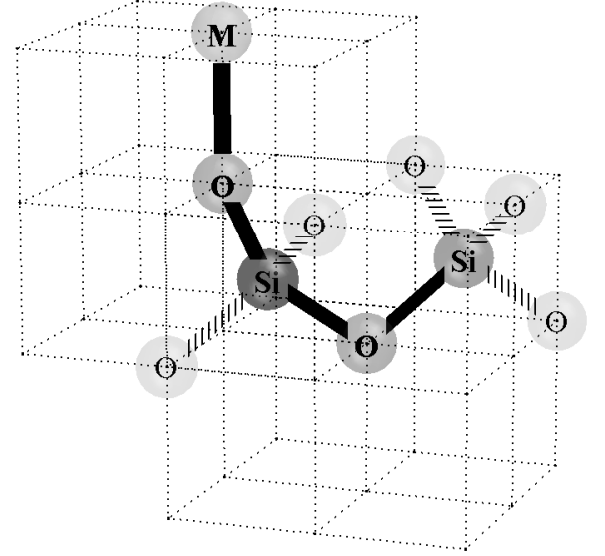


FIG. 7: A face-diagonal configuration M--O--Si--O--Si with overlapping coordination entities Si--O--M and Si--O--Si which share a Si atom. Pale shadings represent the other parts of the silica tetrahedra.

A. Energy parameters

The energetics of the $\text{SiO}_2\text{--M}_2\text{O}$ silicate model is specified by $d = 383$ interaction energies $\mathcal{U}_1, \dots, \mathcal{U}_d$ associated with representatives of the appropriate face-diagonal classes. This is a large number of parameters.

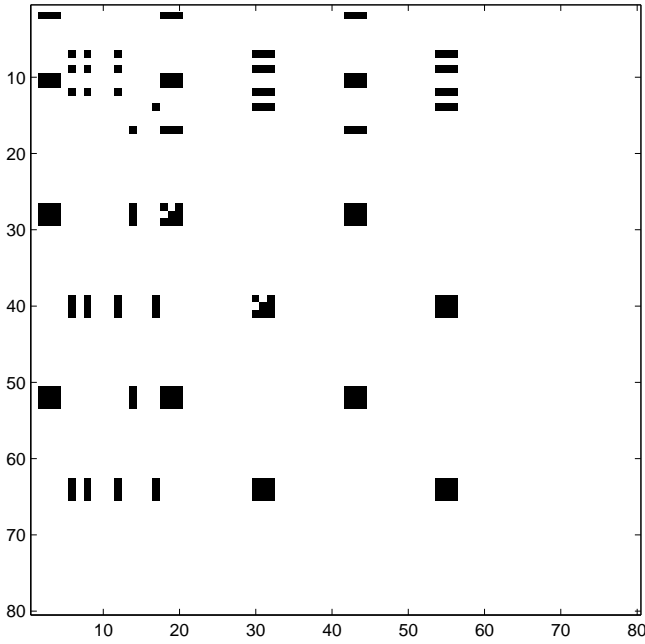


FIG. 8: The black areas represent the 184 face-diagonal configurations which describe overlapping coordination entities associated with the end sites of the $(1, 2, \setminus)$ th face diagonal.

A substantially more economical parameterization is provided by the single site energy model described by (72) whose isotropic version is given by

$$\mathcal{U}_i = \frac{\mathcal{U}_j^{\text{site}} + \mathcal{U}_k^{\text{site}}}{12}, \quad (u, y) \in \Omega_i^{\text{diag}}, u \in \Omega_j^{\text{site}}, y \in \Omega_k^{\text{site}}, \quad (126)$$

where $s = 6$ energies $\mathcal{U}_1^{\text{site}}, \dots, \mathcal{U}_s^{\text{site}}$ are ascribed to representatives of occupied site classes. The latter is similar to the existing thermodynamic models of liquid silicates [15, 16], including the Quasi-Chemical Model and its modifications [17–19], where the internal energy is assumed to be composed of SNNB energies.

A compromise between these two extremes, 383 and 6 energy parameters, consists in applying the single site energy model (126) to nonoverlapping face-diagonal configurations (u, y) , whilst ascribing different energies to representatives of the 31 isotropy equivalence class of overlapping configurations, which gives $6 + 31 = 37$ energy parameters.

B. Constrained minimization

With $\mathcal{N}_{\text{SiO}_2}$ and $\mathcal{N}_{\text{M}_2\text{O}}$ denoting the mole numbers of SiO_2 and M_2O , and

$$x := \frac{\mathcal{N}_{\text{SiO}_2}}{\mathcal{N}_{\text{SiO}_2} + \mathcal{N}_{\text{M}_2\text{O}}} \quad (127)$$

denoting the SiO_2 mole fraction, the centralised mole fractions of Si and M atoms can be calculated by using (10) as

$$y_2 = \frac{\mathcal{N}_{\text{Si}}}{\mathcal{N}_{\text{O}}} = \frac{x}{1+x}, \quad (128)$$

$$y_3 = \frac{\mathcal{N}_{\text{M}}}{\mathcal{N}_{\text{O}}} = \frac{2(1-x)}{1+x} \quad (129)$$

and satisfy the identity $2y_2 + \frac{1}{2}y_3 = 1$. In accordance with Table I, the matrix Υ , given by (100), takes the form

$$\Upsilon = \begin{bmatrix} 12 & 4 & 24 & 24 & 12 & 3 \\ 6 & 2 & 6 & 6 & 0 & 0 \\ 0 & 0 & 24 & 24 & 24 & 6 \end{bmatrix}$$

and has rank two since its rows are linearly related as $\Upsilon_{1\bullet} = 2\Upsilon_{2\bullet} + \frac{1}{2}\Upsilon_{3\bullet}$.

The matrix \widetilde{M} in (103) turns out to be of size 28×383 and full row rank. Therefore, the marginalization and balance equations (106) and (101) lead to $28+2 = 30$ independent linear equations for $383 + 6 = 389$ variables. The fact that the combined set of marginalization and balance equations is highly under-determined favours the Newton scheme of Section XVB for this application which involves 30 Lagrange multipliers instead of 389 dependent variables. However, the above mentioned issues of finding a good initial approximation for the iterative algorithm and avoiding false extrema (which arise from the non-convexity of the problem) require additional investigation.

C. Gibbs energy of mixing

In view of (127), the mixture of $0 < x < 1$ moles of SiO_2 and $1-x$ moles of M_2O contains $1+x$ moles of oxygen atoms. Therefore, the approximation $\widehat{\mathbf{G}}_{\text{face}}(x)$ of the Gibbs energy $\mathbf{G}(x)$ for the liquid silicate can be calculated in terms of (112) as

$$\frac{\widehat{\mathbf{G}}_{\text{face}}(x)}{RT} = (1+x)\widehat{\mathbf{g}}(x),$$

where $R = 8.314 \frac{\text{J}}{\text{mole} \times \text{K}}$ is the universal gas constant, and we have indicated the dependence of $\widehat{\mathbf{g}}$ on the SiO_2 mole fraction x which enters the constrained minimization problem through the balance equations. Hence, the corresponding approximation $\widehat{\Delta \mathbf{G}}_{\text{face}}(x)$ of the Gibbs energy of mixing can be found from

$$\frac{\widehat{\Delta \mathbf{G}}_{\text{face}}(x)}{RT} = (1+x)\widehat{\mathbf{g}}(x) - 2x\widehat{\mathbf{g}}(1) - (1-x)\widehat{\mathbf{g}}(0).$$

D. Mass density

In the framework of the face-to-diagonal reduction of the Kramers-Wannier approximation, the mass density ρ of the

SiO₂-M₂O liquid silicate with chemical composition quantified by (127) is estimated as

$$\begin{aligned}\hat{\rho}_{\text{face}} &= \frac{m_{\text{O}}\mathcal{N}_{\text{O}} + m_{\text{Si}}\mathcal{N}_{\text{Si}} + m_{\text{M}}\mathcal{N}_{\text{M}}}{V_1\mathcal{N}} \\ &= (1 - \hat{S}_0) \frac{(1+x)m_{\text{O}} + xm_{\text{Si}} + 2(1-x)m_{\text{M}}}{(1+x)V_1},\end{aligned}$$

where m_{O} , m_{Si} and m_{M} are the masses of O, Si and M atoms, respectively, and use is made of (8), (128) and (129). As a result of the blended minimization in (112), the fraction \hat{S}_0 of oxygen-free sites of the carrier lattice depends not only on the pressure and temperature through the quantity K defined by (109) but also on the SiO₂ mole fraction x . This dependence reflects a subtle interplay between the geometric constraints and energetics of the system.

XVII. CONCLUSION

For the class of NNE-constrained interacting particle systems on the simple cubic lattice, which is relevant for statistical mechanical modeling of disordered condensed media, we have carried out a face-to-diagonal reduction of the Kramers-Wannier entropy density approximation. The latter represents the second level in the hierarchy of CVM approximations and takes into account the statistical correlations of the equilibrium state within faces of cubic cells.

Using a separation-of-variables technique in the framework of this entropy density approximation, we have obtained equations for approximate computation of the equilibrium Gibbs free energy in a class of lattice models of chemical systems with complex site configurations and short-range ordering, and a three-loop architecture for its numerical implementation has been proposed.

We have outlined an application of the statistical mechanical approach to thermodynamic modeling of a binary liquid silicate formed from silica and the oxide of a univalent metal. The combinatorial and numeric aspects of the computer implementation of the model have been discussed. The results on model calibration for specific systems (such as sodium silicate SiO₂-Na₂O) are postponed in view of additional work on a reliable algorithm required for solving the subsidiary non-convex minimization problem.

The NNE-constrained cubic lattice setting is amenable to further refinement in the form of a “quasi-tetrahedron” reduction of Kikuchi’s cube approximation of CVM for more subtle predictions of thermodynamic properties of such systems. The latter development will be communicated in subsequent publications.

Acknowledgments

This work was carried out in 2007–2008 while the author was with the University of Queensland. The work was supported by a UQ Research and Development grant and the Australian Research Council. Helpful discussions on chemical

thermodynamics with Dmitry Saulov, Alex Kondratiev, Eugene Jak, Peter Hayes and Alex Klimenko are also gratefully acknowledged.

Appendix A: Unique solvability of the decic equation

The decic equation (119) is equivalent to χ being a zero of the function

$$f(x) := x + \frac{1}{2} \left(e^{-K/3} g(\theta x) - 1 \right)$$

on the interval $(0, 1/2]$. Here,

$$g(x) := (1 - 2x)^3 (1 - x)^{-7/3}$$

is a strictly convex function which decreases strictly from $g(0) = 1$ to $g(1/2) = 0$. The monotonicity of g follows from

$$g'(x) = \frac{1}{3} (1 - 2x)^2 (1 - x)^{-10/3} (4x - 11) < 0.$$

Hence, for any $K > 0$ and $\theta \in [1/2, 1)$, the function f satisfies

$$f(0) = \frac{e^{-K/3} - 1}{2} < 0 < \frac{e^{-K/3} g(\theta/2)}{2} = f(1/2),$$

which, by the Intermediate Value Theorem, implies that the equation $f(x) = 0$ has a root on $(0, 1/2]$. The uniqueness of the root follows from the convexity of f inherited from g . The latter is established by

$$g''(x) = \frac{2}{9} (1 - 2x)(1 - x)^{-13/3} (-4x^2 + 22x + 17) \geq 0,$$

where both roots of the rightmost quadratic polynomial are beyond the interval $(0, 1/2]$.

Appendix B: Newton iterations

The constrained optimization problem of Section XV B is a particular case of the minimization problem

$$f(p) := \langle a, p \rangle + \langle b, \Lambda(p) \rangle \longrightarrow \min,$$

where the minimum is taken over a finite-dimensional column-vector p of probabilities subject to the system of linear constraints

$$Ap = q.$$

Here, the function Λ , defined by (17), applies entry-wise; a , b and q are appropriately dimensioned column-vectors, with b consisting of nonzero entries, and the matrix A is of full row rank. The linear part $\langle a, p \rangle$ of the objective function is the internal energy term, while $\langle b, \Lambda(p) \rangle$ originates from an entropy cumulant expression, typical for CVM. If the entries of

b are all positive, then the function f is strictly convex. However, if some of them are negative, as is the case in the above mentioned problem of Section XV B, the overall convexity of f under the linear constraints is a nontrivial issue [11] which complicates the numerical solution of this problem. At the level of first-order necessary conditions of optimality for relatively interior points, the stationarity of the Lagrange function

$$\mathcal{L}(p, \tau) := f(p) - \langle \tau, Ap - q \rangle$$

with respect to p is equivalent to that the gradient of f satisfies

$$f'(p) = A^T \tau,$$

which allows p to be expressed in terms of the column-vector τ of Lagrange multipliers as

$$\hat{p} := \exp((A^T \tau - a)/b - \mathbf{1}).$$

Here, the exponential function and division are applied to vectors entry-wise, and $(\cdot)^T$ denotes the matrix transpose. The Jacobian matrix of the map \hat{p} with respect to τ is computed as

$$\hat{p}' = (f''(\hat{p}))^{-1} A^T = \text{diag}(\hat{p}/b) A^T,$$

where

$$f''(p) = \text{diag}(b/p)$$

is the Hessian matrix of the function f , and $\text{diag}(v)$ is the diagonal matrix with the vector v over the main diagonal. Hence, in order to satisfy the linear constraints, the Newton iterations for τ take the form

$$\tau \mapsto \tau - (A \text{diag}(\hat{p}/b) A^T)^{-1} (A \hat{p} - q).$$

Here, the invertibility of the matrix is not guaranteed if b contains both positive and negative entries, even though A is of full row rank. By assuming that the iterates converge to $\hat{\tau}$ and using the identity

$$\langle p, f'(p) \rangle = f(p) + \langle b, p \rangle,$$

it follows that the corresponding value of the function is calculated as

$$f(\hat{p}) = \langle q, \hat{\tau} \rangle - \langle b, \hat{p} \rangle.$$

In the complicated case, where some entries of the vector b are negative, a sufficient condition for the point \hat{p} to be a local minimum under the linear constraints is the positive definiteness of the Hessian matrix $f''(\hat{p})$ on the null space of the matrix A .

-
- [1] I. Vladimirov and E. Jak, *J. Chem. Phys.* **126**, 164502 (2007).
 - [2] A. Verhagen, *J. Chem. Phys.* **67**, 5060 (1977).
 - [3] R. Baxter, *Annals of Combinatorics* **3**, 191 (1999).
 - [4] K. Huang, *Statistical Mechanics* (John Wiley & Sons, New York, 1987), 2nd ed.
 - [5] G. An, *J. Statist. Phys.* **52**, 727 (1988).
 - [6] R. Kikuchi, *Phys. Rev.* **81**, 988 (1951).
 - [7] R. Kikuchi, *J. Chem. Phys.* **19**, 1230 (1951).
 - [8] M. Kurata, R. Kikuchi, and T. Watari, *J. Chem. Phys.* **21**, 434 (1953).
 - [9] R. Kikuchi and K. Masuda-Jindo, *Calphad* **26**, 33 (2002).
 - [10] J. Moran-Lopez and J. Sanchez, eds., *Theory and Applications of the Cluster Variation and Path Probability Methods* (Plenum Press, New York, 1996).
 - [11] A. Pelizzola, *J. Phys. A: Math. Gen.* **38**, R309 (2005).
 - [12] T. Cover and J. Thomas, *Elements of Information Theory* (Wiley, Hoboken, New Jersey, 2006), 2nd ed.
 - [13] F. Cotton, *Chemical Applications of Group Theory* (Wiley, New York, 1990), 3rd ed.
 - [14] H. Enderton, *A Mathematical Introduction to Logic* (Harcourt Academic Press, San Diego, 2001), 2nd ed.
 - [15] D. Gaskell, *Canad. Met. Quaterly* **20**, 3 (1981).
 - [16] B. Mysen, *Structure and Properties of Silicate Melts* (Elsevier Science Publishers, Amsterdam, 1988).
 - [17] M. Blander and A. D. Pelton, in *Proc. 2nd Int. Symp. Met. Slags and Fluxes* (TMS-AIME, Warrendale, PA, 1984), pp. 295–304.
 - [18] A. Pelton and M. Blander, *Met. Trans. B* **17B**, 805 (1986).
 - [19] A. Pelton, S. Degterov, G. Eriksson, C. Robelin, and Y. Dessureault, *Met. Mat. Trans. B* **31B**, 651 (2000).
 - [20] R. Kindermann and J. Snell, *Markov Random Fields and Their Applications* (American Mathematical Society, Providence, 1980).
 - [21] C. Preston, *Random Fields* (Springer-Verlag, Berlin, 1976).
 - [22] J. Conway and N. Sloane, *Sphere Packings, Lattices and Groups* (Springer-Verlag, New York, 1988).
 - [23] A. Shiryaev, *Probability* (Springer, Berlin, 1995), 2nd ed.
 - [24] In general, the joint probability distribution of two or more random variables is not specified uniquely by their marginal distributions.